

Updated January 2007
**Chesapeake Bay Program Guidance and Policies for Data,
Information and Document Deliverables Submission**

The *Chesapeake Bay Program Guidance and Policies for Data, Information and Document Deliverables Submission* describes the guidelines and policies governing the submission of electronic deliverables to the Chesapeake Bay Program. Electronic deliverables can be submitted directly to the Chesapeake Bay Program Office or served on the Internet as part of the Chesapeake Information Management System (CIMS).

The full guidance document "Chesapeake Bay Program Guidance for Data Management" is available in electronic format on the web at www.chesapeakebay.net/cims. Below are excerpts from that guidance.

Applicability

These guidelines and policies must be followed by all agencies, institutions, and organizations participating in data and information collection, processing, document generation and submittal to the Chesapeake Bay Program under grant or cooperative agreement funding. The Chesapeake Bay Program has adopted these guidelines and policies in order to improve coordination, compatibility, standardization, and information access across all the Bay Program partners. In addition to these guidelines and policies, any activities funded with Federal Government funds, must also adhere to applicable Federal Information Processing Standards (FIPS) (www.itl.nist.gov/div897/pubs/)

CBP Data/information Management and Document Deliverables Guidelines and Policies

The Chesapeake Bay Program has adopted the following guidelines and policies pertaining to data and information collection, processing, document generation and submittal to the Chesapeake Bay Program under grant or cooperative agreement funding. Any deviations from these guidelines and policies must be documented in the work plan and approved by the EPA Project Officer.

- Data, Information, and Document Deliverables Requirements
- Deliverable Serving vs. Submission Policy
- Locational Data Policy
- Map Coordinate Datum policy
- Map Coordinate Projection Guideline
- Metadata Policy
- Common Station Names Guideline
- Common Data Dictionary Guideline
- Common Database Design Guideline
- Calendar Date Policy
- Common Method Codes Guideline
- Data Reporting Guideline
- ITIS Biological Nomenclature Policy

Data, Information, and Document Deliverables Requirements

Recipients are required to submit data, information, and document deliverables in electronic format unless exceptions are specified in the grant or cooperative agreement work plan. Electronic deliverables include but are not limited to reports, graphics, spreadsheets, imagery, data files, audio, and digital video products.

All data, information, and documents funded by the Chesapeake Bay Program whether through direct Chesapeake Bay Program funding or indirect matching funds are public information and shall be made available to the public unless there is a grant/cooperative agreement award condition that specifies otherwise. In addition, source data collected and processed in the creation of a deliverable should also be submitted when practical. If source data is submitted, it should also be delivered in electronic format. All deliverables must have associated metadata.

Document Type	Acceptable Formats
Text	Preferred: Microsoft Word (DOC) Portable Document Format (PDF) * ASCII Text Extensible Markup Language (XML)
Spreadsheet	Preferred: Microsoft Excel 97 or higher With Prior Approval: PageMaker Lotus 1-2-3 QuatroPro Tab/Comma delimited text files
Database	Preferred: Microsoft Access 97 or higher Microsoft SQL Server Extensible Markup Language (XML)
Graphics	Preferred: TIFF GIF JPEG SVG
Geographic Information System	Preferred: ESRI Spatial Database Engine (SDE) or personal geodatabase ESRI coverages, grids, shapefiles ArcExport non-compressed (E00)

*** Data tables within PDF documents must be delivered in one of the spreadsheet formats.**

Deliverable Serving vs. Submission Policy

Recipients can submit deliverables directly to the Chesapeake Bay Program or serve deliverables from a data/web server. The preferred method for serving data is through an extensible markup language (XML) web service. Recipients who plan to directly serve their grant/cooperative agreement deliverables through their own data server/web site must have signed a CIMS Memorandum of Understanding with the Chesapeake Bay Program.

Locational Data Policy

The Chesapeake Bay Program adheres to the EPA's locational data policy, which requires consistent use of latitude/longitude coordinates to identify the location of entities. All data, containing spatial and/or specific geographic locations, collected or assembled under a Chesapeake Bay Program grant or cooperative agreement or to be served on the Internet via the Chesapeake Information Management System, must have latitude and longitude information for each entity. Projects not creating or reporting spatial data, but-confined to a given project location(s), shall include the latitude/longitude of the location(s) within the study/final report.

In accordance with Chesapeake Bay Program locational data policy, the recipients agree to ensure that latitude and longitude coordinates (given in degrees and decimal degrees) are provided for all sites for which data are collected and accurate to the level required for the purpose of the application of the data. Field measured locations shall be accurate to the best practical geographic positioning method. Currently, Differential Global Positioning System (GPS) equipment can reliably provide coordinates accurate to within 10-25 meters (5 decimal places in decimal degrees), and is the preferred method of point location determination. Applications such as station monitoring locations should provide locational data with accuracy to that level. Other applications, such as digitizing points or watershed boundaries from Mylar media maps, can not provide accuracy better than that of the original map, and can not match the accuracy of GPS or surveyed locations. Remote sensing platforms can now collect sub-meter resolution data (6 decimal places in decimal degrees). Therefore, it is required that metadata be provided for all data and must include a measurement of the accuracy of the coordinates and the original source material and methods for obtaining the coordinates. It is the responsibility of data generators/providers to provide coordinates accurate to the level that is practical for the intended application, and to document the accuracy of those coordinates. The recipient further agrees to document, in writing, that locational data were derived using an approved method and recorded in accordance with federal regulations and other EPA requirements, noted in the "Authorities" section of the EPA's policy. Recipient shall include in their work plan an assurance to comply with this requirement.

Map Coordinate Datum Policy

The Chesapeake Bay Program has adopted the policy that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS shall utilize either the North American Datum 1983 (NAD83) or World Geodetic System 1984 (WGS84) horizontal reference and the North American Vertical Datum 1988 (NAVD88) vertical reference. Most likely, organizations have been using NAD27 horizontal reference since USGS maps were

historically created using this reference. The requirement to use NAD83 or WGS84 will require conversion of latitudes and longitudes using NAD27 to NAD83/WGS84. Metadata reporting requires specification of the horizontal and vertical datum where applicable.

Map Coordinate Projection Guideline

The Chesapeake Bay Program has adopted the policy that the standard projection for geographic information system (GIS) files maintained at the Chesapeake Bay Program Office (CBPO) shall be UTM Zone 18 (meters) for all data within the Chesapeake Bay basin. For larger or national GIS data files, the standard projection for GIS files maintained at the CBPO shall be Albers Conical Equal Area (meters). This policy was established to provide consistency in computing distance and area calculations, map shapes, and to facilitate database design and maintenance, and based on the recommendation of USGS. GIS and data files containing spatial data, must have coordinates reported as latitude and longitude (decimal degrees) as per the Locational Data Policy. Ideally, it is requested that information containing projected coordinates, also report coordinates in UTM Zone 18. Partner organizations that have historically maintained GIS files in another projection or coordinate system are exempt from this policy (unless they are developing or providing data products as part of a Chesapeake Bay Program initiative) since the effort to convert large historical holdings would be prohibitive.

Metadata Policy

The Chesapeake Bay Program has adopted the policy, consistent with Presidential Executive Order # 12906, that all data generated or collected using federal funds, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS, shall be accompanied by metadata (descriptive information about the data, often referred to as documentation) that fully conforms to the Federal Geographic Data Committee's requirements for metadata. Metadata created for Chesapeake Bay Program shall also be delivered to the EPA or other federal Clearinghouse as a requirement to fulfilling this policy and related grant or contract conditions. The FGDC guide for creating metadata is the *Content Standard for Digital Geospatial Metadata Workbook* (www.fgdc.gov/metadata).

The Chesapeake Bay Program has also adopted the policy, that all data generated or collected using federal funds, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS, shall adhere to the National Biological Information Infrastructure's (NBII) Metadata Standard, where applicable. The NBII Metadata Standard, popular for environmental programs, provides extensions to the FGDC Metadata for documenting biological data and information. The NBII Biological Data Profile can be found at: www.nbii.gov.

Data to be accessed on the Internet via CIMS must follow the *CIMS Metadata Reporting Guidelines* established by the Chesapeake Bay Program. This Guideline was established to facilitate entering consistent, accurate metadata to ensure the information about the Chesapeake Bay will be easily available, and used appropriately. The *CIMS Metadata Reporting Guidelines* is also accessible on the CIMS Internet Web Page. The COMET system (www.chesapeakebay.net/comet) provides a streamlined, easy to use tool for entering metadata that meets CIMS and FGDC requirements.

Common Station Names Guideline

The Chesapeake Bay Program has adopted the guideline that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS should utilize a consistent set of common station names for identifying and reporting monitoring station locations. It is the data provider's responsibility to comply with this guideline. The purpose of this guideline is to create one master table of station names, to the extent possible, to reduce confusion among cooperating agencies. The Station Names table, maintained on the Chesapeake Bay Program web site, should serve as the master list. Updates to this table that are required by data submitters shall be coordinated with the Information Management Subcommittee and/or the Monitoring and Assessment Subcommittee to maintain one consistent stations names list.

Common Data Dictionary Guideline

The Chesapeake Bay Program has adopted the guideline that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS should utilize the CBP common data dictionary for defining all data elements and units of measure. It is the, data provider's responsibility to comply with this policy. The purpose of this guideline is to create one data dictionary, to the extent possible, to reduce confusion among cooperating agencies. Updates required by data submitters to the dictionary shall be coordinated with the CIMS Workgroup to maintain one consistent data dictionary.

Common Database Design Guideline

The Chesapeake Bay Program has adopted the guideline that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS should utilize the CBP common database design for managing data. It is the data provider's responsibility to comply with this guideline. The purpose of this guideline is to use common database designs, to the extent possible, to simplify data formatting and sharing. Modifications to the common database design shall be coordinated with the CIMS Workgroup to maintain consistency in the database structure. If the Chesapeake Bay Program agencies do not have a pre-defined database that is acceptable for the work being conducted, the grantee/contractor should work with the funding agency to develop a database design that suits the requirements of the work. The database design should maintain maximum compatibility with other Chesapeake Bay Program database designs.

Calendar Date Policy

The Chesapeake Bay Program has adopted the standard that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS should adhere to the Federal Information Processing Standard, Representation for Calendar Date and Ordinal Date for Information Interchange (FIPS PUB 4- 1).

This standard states, "For purposes of electronic data interchange in any recorded form among U.S. Government agencies, National Institute of Standards and Technology (NIST) highly

recommends that four-digit year elements be used". The year should encompass a two-digit century that precedes, and is contiguous with, a two-digit year-of-century (e.g., 1999, 2000, etc.). In addition, optional two-digit year time elements specified in ANSI X3.30-1985(RI991) should not be used for the purposes of any data interchange among U.S. Government agencies. Therefore, it is required to report and store all dates using four digits for the year. In addition to facilitating data sharing, this requirement reduces the complications of processing date data after the millennium rollover at year 2000.

Common Method Codes Guideline

The Chesapeake Bay Program has adopted the guideline that all data generated or collected for, submitted to the CBP, or served on the Internet via CIMS should utilize the CBP Method Code tables. The purpose of this guideline is to use standardized method codes, to the extent possible, to simplify data coding and sharing. The methods used by monitoring agencies and analytical laboratories are critical in providing accurate measurements. Knowing the field and laboratory methods used is critical; therefore capturing the methods is a high priority during database development. **Existing method codes are listed online in the respective CIMS online data dictionary at <http://www.chesapeakebay.net/data/index.htm>. It is the data provider's responsibility to comply with this guideline in consultation with the appropriate CBP Database Manager.** If the CBP lacks pre-defined method codes that are acceptable for the work being conducted, the grantee/contractor should work with **the CBP Database Manager** to develop method codes that suit the requirements of the work, while maintaining maximum compatibility with other CBP codes.

Numeric Data Reporting Guideline

The Chesapeake Bay Program has adopted the guideline that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS should report numeric data elements at the same level of precision as that of the original measurement. The exact precision of recorded values must be maintained. This guideline has a significant impact on data analysis and the decisions made based on these analyses.

Values should not be zero-filled to greater precision than actually recorded. For instance, if the measured value is 0.03, then the reported value should be 0.03 @ and not 0.030, which would imply precision to the third decimal place. For values that are recorded as below detection, a detection flag (in a separate data field) shall be used to identify the value as below (**i.e., "<"**) the detection limit of the method, and the value shall be reported as the detectable limit. **The funding agency may request the reporting of instrument readings below the method detection limits, including negative readings.** In this case, zero values shall be reported only if the measured or recorded value is zero.

Values that are missing shall be reported as missing or null or nil, to identify values that were sampled but no value was obtained. Missing, null, or nil values are different than those that were never sampled, which should be recorded as a blank field, if they are recorded at all. It is the responsibility of the data submitter to record in the metadata, how measurements are coded, as well as the accuracy of the measurements.

It is important to note that some software tools used in data processing may represent the data internally with more precision than the original measurement, and/or may round the value. For instance even though a value of 0.3 was entered, the value may be stored and reported as 0.299999.

ITIS Biological Nomenclature Policy

The Chesapeake Bay Program has adopted the policy that all data generated or collected for, submitted to the Chesapeake Bay Program, or served on the Internet via CIMS should utilize the ITIS (www.itis.usda.gov/) biological names for identifying and reporting species. It is the data provider's responsibility to comply with this policy. The purpose of this policy is to create one master table of species names, to the extent possible, to reduce confusion among cooperating agencies. The ITIS taxonomy table, maintained on the ITIS web site, should serve as the master list. Updates to this table that are required by data submitters shall be coordinated with the CIMS Workgroup to maintain one consistent species name list.